



Plant Archives

Journal homepage: <http://www.plantarchives.org>

DOI Url : <https://doi.org/10.51470/PLANTARCHIVES.2025.v25.no.2.196>

ASSESSMENT OF GENETIC DIVERSITY USING PCA FOR YIELD ATTRIBUTING AND GRAIN QUALITY TRAITS IN AROMATIC SHORT-GRAIN RICE (*ORYZA SATIVA* L.) MUTANTS ALONG WITH THEIR PARENTAL LINES

Subhra Acharyya*, Shubhashree Panda and Deepak Sharma

Department Genetics and Plant Breeding, College of Agriculture, IGKV, Raipur, Chhattisgarh, India

*Corresponding author E-mail: subhraacharyya05@gmail.com

(Date of Receiving-06-06-2025; Date of Acceptance-12-08-2025)

ABSTRACT

A total of 52 aromatic short grain rice genotypes were investigated for yield attributing traits and grain quality traits using principal component analysis (PCA) for the determination of variation pattern, the relationship among genotypes and its traits. Five main components with eigenvalues greater than 1.0 were identified by PCA; these components explained 81.16% of the variation in yield traits and 84.46% of the variation in grain quality traits, respectively. PC1 was mostly affected by plant height, spikelet fertility and filled grains; PC2 was affected by total and effective tillers; PC3 was affected by hundred seed weight and grain yield; PC4 was affected by days to 50% flowering and unfilled grains; and PC5 was affected by spikelet fertility. Kernel L/B ratio and kernel length ruled PC1 for quality features; milling percentage, amylose content and head rice recovery governed PC2; elongation ratio and cooked rice width governed PC3; cooked rice dimensions governed PC4; and paddy and kernel width governed PC5.

Key words: principal component analysis, aromatic short grain rice genotypes, genetic diversity.

Introduction

Rice (*Oryza sativa* L.) is the world's most popular staple food crop, feeding more than half of the world's population and accounting for approximately 21% of global per capita energy and 15% of protein intake (IRRI, 2023). Rice was grown on around 170 million hectares worldwide and produced 543.6 million tons of milled rice in 2024-25, making it the foundation of global food security (FAO, 2024). With more than 90% of the world's rice production, Asia is the dominant region in rice growing. With an expected 51.42 million hectares under cultivation and an output of 150 million tons of raw rice in 2024-2025, India is the world's second-largest producer of rice, after China (Ministry of Agriculture and Farmers Welfare, 2025). Rice continues to be essential to rural livelihoods and food security, accounting for about 43% of India's total food grain production (AgriStat, 2024). Rice is grown in a variety of agroecological environments, including irrigated, rainfed lowland, upland and flood-prone

locations, making it extremely flexible. However, the increasing global population, shifting food tastes and climate unpredictability necessitate improvements not only in production but also in grain quality features such as cooking behaviour, texture, nutritional content and aromas (Sohgaura *et al.*, 2015; Koutroubas *et al.*, 2004).

Aromatic and specialized rice cultivars are becoming increasingly popular due to their high market value and consumer preference. These include classic basmati varieties and local short-grain aromatic landraces with distinct aroma and grain quality, which have economic and cultural significance (Yumnam *et al.*, 2015; Ghareyazie *et al.*, 1997).

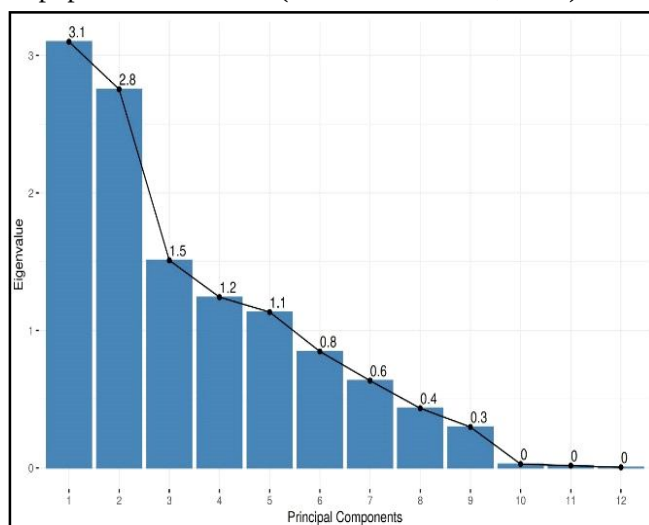
Therefore, rice genetic improvement remains critical to addressing the challenges of yield stability, grain quality enhancement and sustainability. Evaluation and use of genetic variation in rice germplasm are critical methods in modern breeding programs for ensuring food and nutritional security (Vanja and Babu, 2006; Nachimuthu

Table 1: List of rice genotypes used in this study.

S.	Genotype	S.	Genotype	S.	Genotype	S.	Genotype
1	Gopalbhog Mutant-4	14	Vishnubhog Parent(fine)	27	Laxmibhog Mutant-21	40	Laxmibhog Parent
2	Gopalbhog parent	15	Kalanamak Mutant-2	28	Laxmibhog Parent	41	Ambemohar Mutant-1
3	Chinnor Mutant-5	16	Kalanamak Parent	29	Ambemohar Mutant-1	42	Ambemohar Parent
5	Chinnor Parent	17	Ramjeera Mutant-2	30	Ambemohar Parent	43	Laxmibhog Mutant-21
5	Shyamjeera Mutant-5	18	Ramjeera Parent	31	Laxmibhog Mutant-21	44	Laxmibhog Parent
6	Shyamjeera Parent	19	C.G. Jawaphool Trombay	32	Laxmibhog Parent	45	Ambemohar Mutant-1
7	Nagri Dubraj Mutant-5	20	Jawaphool Parent	33	Ambemohar Mutant-1	46	Tilkasturi Parent
8	Nagri Dubraj Parent	21	TCDM-1 (Check)	34	Ambemohar Parent	47	Badshabhog Mutant-2
9	Jeeraphool Mutant-2	22	Mai Dubraj Parent	35	Laxmibhog Mutant-21	48	Badshabhog Parent
10	Jeeraphool Parent	23	Kadamphool Mutant-1	36	Laxmibhog Parent	49	Tulsimanjiri Mutant-6
11	Loktimachi Mutant-9	24	Kadamphool Parent	37	Ambemohar Mutant-1	50	Tulsimanjiri Parent
12	Loktimachi Parent	25	Dhania Dhan Mutant-1	38	Ambemohar Parent	51	Loktimachi Mutant-29
13	TCVM (Check)	26	Dhania Dhan Parent	39	Laxmibhog Mutant-21	52	Grassy Dubraj

et al., 2014). Principal Component Analysis (PCA) is a strong multivariate statistical tool that is widely used in plant breeding programs to evaluate genetic diversity and structure within germplasm collections. It reduces data dimensionality while keeping maximum variability by converting correlated variables into a set of uncorrelated components known as principal components (Anderson, 1972). These components capture the major patterns of variance, allowing researchers to more effectively evaluate complex datasets.

PCA has been frequently used in rice (*Oryza sativa* L.) to find the most discriminating traits that contribute to genetic divergence among genotypes, particularly yield and grain quality parameters (Maji and Shaibu, 2012; Ramakrishnan *et al.*, 2016). By summarizing trait variability, PCA allows researchers to remove duplication in datasets caused by inter-trait correlations and extract the fewest number of components that explain the majority of population variation (Nachimuthu *et al.*, 2014).

**Fig. 1:** Scree plot diagram using principal components of rice genotypes for yield attributing traits.

The use of PCA, particularly when combined with cluster analysis, improves the resolution of genetic diversity estimates and aids in grouping genetically diverse genotypes, which is critical for maximizing genetic gains in aromatic rice breeding programs (Maji *et al.*, 2012; Sivaranjani *et al.*, 2010).

Materials and Methods

The study used fifty-two aromatic short grain rice mutants and their parents (Table 1). During the *kharif* season of 2024, the experiment was carried out in the Department of Genetics and Plant Breeding's Instructional Cum Research Farm, College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, India. The experiment used a Randomized Complete Block Design (RCBD), which was replicated twice. For each replication, five plants were chosen at random for observation.

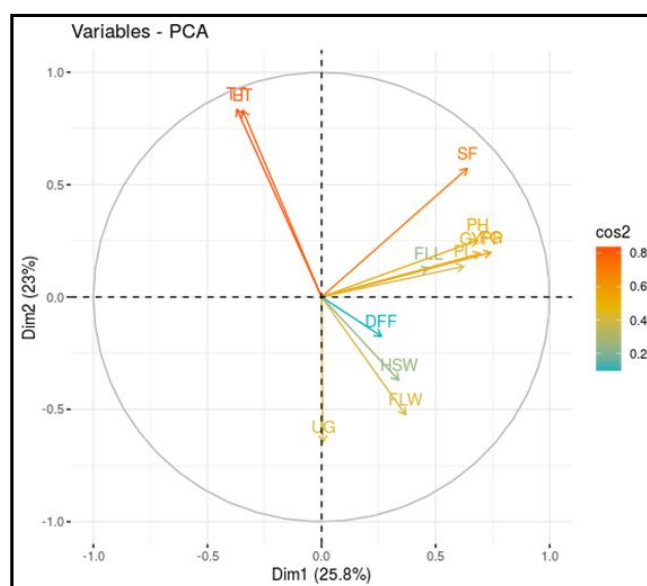
**Fig. 2:** Biplot of 12 yield attributing traits across PC1 and PC2 for yield attributing traits.

Table 2: Eigen values, Percentage of variation and Cumulative percentage for principal components for yield attributing traits.

Principle components	Eigen value	% of variance	Cumulative % of variance
PC1	3.1	25.835	25.835
PC2	2.754	22.952	48.787
PC3	1.51	12.583	61.37
PC4	1.242	10.351	71.721
PC5	1.133	9.444	81.165

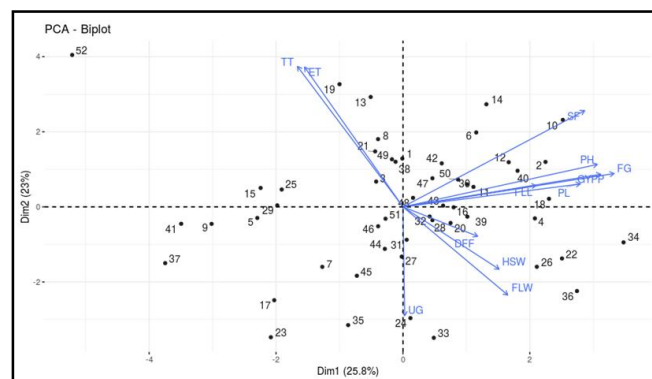
Days to 50% flowering, plant height (cm), panicle length (cm), number of productive tillers per plant, flag leaf length (cm), flag leaf width (cm), number of filled grains per panicle, number of unfilled grains per panicle, hundred seed weight (g), grain yield per plant (g) were the observations that contributed to the yield.

Additionally, the following characteristics of grain quality were noted: elongation ratio, alkali spreading value, amylose content percentage, hulling percentage, milling percentage, head rice recovery percentage, paddy length (mm), paddy width (mm), kernel length (mm), kernel width (mm), kernel L/B and cooked rice length (mm). A vernier caliper was used to measure the milled rice's length, milled rice's breadth and paddy's length and width.

The sensory evaluation approach proposed by Sood and Siddiq (1978) likewise revealed the presence of scent. PCA was used to categorize the patterns of variance. As suggested by Jeffers (1967), those PCs having Eigen values larger than one were chosen. We computed correlations between the original qualities and their corresponding Principal Components (PCs). The program R Studio was used to perform statistical analysis. Using the Euclidean distance measure and the agglomerative clustering approach, cluster analysis was carried out.

Results and Discussion

In the present study, PCA was performed using yield and quality related traits in 52 aromatic rice germplasm accessions.

**Fig. 3:** The Biplot of aromatic rice genotypes for PC1 and PC2 for yield attributing traits.**Table 3:** Contribution of five principal components to variation for yield attributing traits.

Variables	PC1	PC2	PC3	PC4	PC5
DFF	0.149	-0.105	-0.019	0.691	-0.098
PH	0.389	0.152	-0.234	-0.102	-0.273
PL	0.354	0.082	-0.382	-0.267	-0.168
TT	-0.211	0.503	0.128	0.162	-0.198
ET	-0.197	0.501	0.15	0.168	-0.185
FLL	0.266	0.077	-0.234	-0.1	-0.473
FLW	0.209	-0.316	0.292	-0.164	0.213
UG	0.004	-0.39	-0.223	0.435	-0.192
FG	0.422	0.12	-0.104	0.324	0.381
SF	0.363	0.345	0.098	-0.061	0.428
HSW	0.192	-0.223	0.572	-0.141	-0.4
GYPP	0.395	0.117	0.472	0.177	-0.14

DFF= Days to 50% flowering, PH= Plant height (cm), PL= Panicle length (cm), TT= Total tillers, ET= Effective tillers, FLL= Flag leaf length (cm), FLW= Flag leaf width (cm), UG= Unfilled grains per panicle, FG= Filled grains per panicle, SF= Spikelet fertility, HSW= Hundred seed weight (g), GYPP= Grain yield per plant (g)

For yield attributing traits out of twelve pcs, five exhibited more than 1.0 eigen values and exhibited 81.165% total variability among the characters studied. Scree plot elucidated the variation percentage between Eigenvalues and the Principal components (Fig. 1). The principal components having more than one eigen value showed more variation among the aromatic rice germplasm accessions for the selection of the diverse parents. Among the five principal components of yield attributing traits, PC1 shared high proportion of total variation 25.835% and the rest of the four principal components viz., PC2, PC3, PC4 and PC5 contributed 22.952%, 12.583%, 10.351% and 9.444% of the total variance respectively. Contribution of twelve yield attributing traits to the principal components is presented

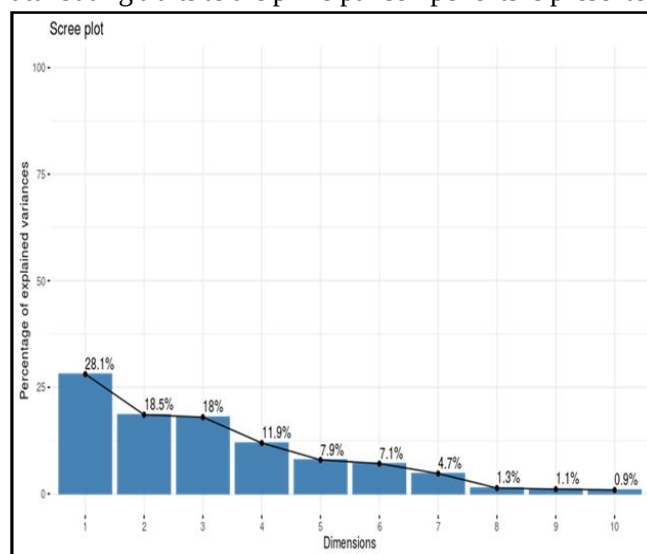
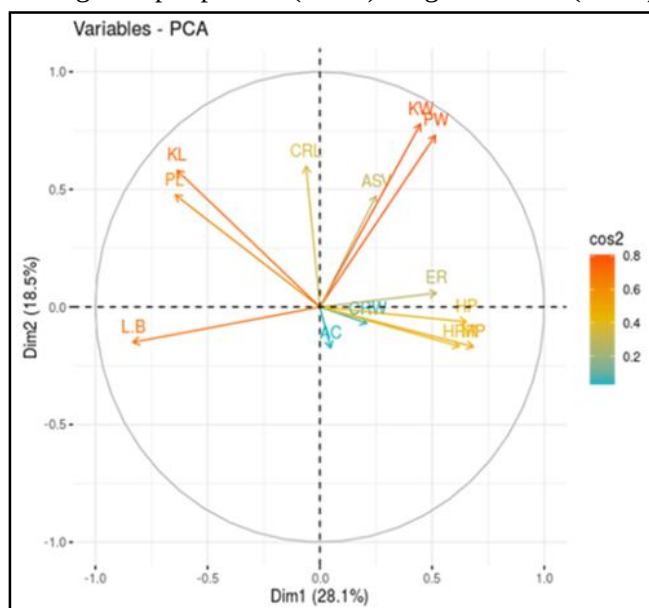
**Fig. 4:** Scree plot diagram using principal components of rice genotypes for grain quality traits.

Table 4: Eigen values, Percentage of variation and Cumulative percentage for principal components for grain quality traits.

Principle components	Eigen value	% of variance	Cumulative % of variance
PC1	3.65	28.078	28.078
PC2	2.41	18.541	46.619
PC3	2.337	17.975	64.594
PC4	1.549	11.913	76.507
PC5	1.032	7.939	84.446

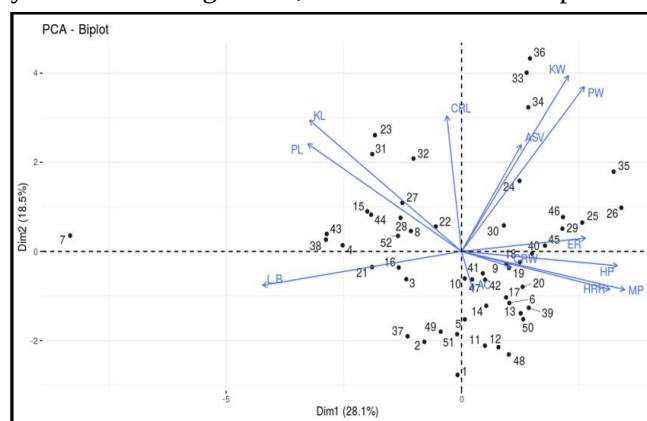
in Table 3. Plant height (0.389), panicle length (0.354), number of filled grains per panicle (0.422), spikelet fertility (0.363), grain yield per plant (0.395) showed maximum positive loading in PC1 while the number of effective tillers (-0.197) and total tillers (-0.211) showed negative loadings. In PC2, the parameters viz., spikelet fertility (0.345), the number of effective tillers (0.501) and total tillers (0.503) showed maximum positive loading and number of unfilled grains per panicle (-0.39) and flag leaf width (-0.316) showed maximum negative loadings. In PC3, traits like hundred seed weight (0.572), grain yield per plant (0.472), flag leaf width (0.292), total tillers (0.128), effective tillers (0.15), spikelet fertility (0.098) showed positive loading whereas, further traits showed negative loadings. These traits are largely engaged in the divergence and they also carry most of the variability. In PC4, the parameters viz., days to 50% flowering (0.691), the number of filled grains per panicle (0.324), the number of unfilled grains per panicle (0.435), grain yield per plant (0.177), total tillers (0.162), effective tillers (0.168) showed positive loading while other characters showed negative loadings. Spikelet fertility (0.428), number of filled grains per panicle (0.381), flag leaf width (0.213)

**Fig. 5:** Biplot of grain quality traits across PC1 and PC2.**Table 5:** Contribution of five principal components to variation for yield attributing traits.

Variables	PC1	PC2	PC3	PC4	PC5
HP	-0.342	0.041	-0.459	0.098	-0.023
MP	-0.358	0.109	-0.445	0.065	-0.047
PdL	0.338	-0.307	-0.329	0.008	-0.001
PW	-0.27	-0.47	0.111	-0.119	0.201
KL	0.333	-0.374	-0.289	0.076	0.034
KW	-0.235	-0.501	0.096	-0.183	0.103
L/B	0.437	0.097	-0.268	0.225	-0.037
CRL	0.032	-0.387	-0.023	0.572	-0.278
CRW	-0.11	0.045	0.117	0.488	0.302
ER	-0.272	-0.038	0.3	0.471	-0.313
ASV	-0.131	-0.304	-0.104	-0.258	-0.337
AC	-0.025	0.113	-0.023	-0.159	-0.737
HRR	-0.326	0.108	-0.436	0.059	0.146

HP=Hulling %, MP= Milling %, PdL=Paddy length (mm), PW= Paddy width(mm), KL= Kernel length(mm), KW=Kernel width(mm), L/B= Kernel length/ Kernel breadth, CRL= Cooked rice length(mm), CRW= Cooked rice width(mm), ER= Elongation ratio, ASV= Alkali Spreading Value, AC= Amylose content %, HRR= Head rice recovery %.

showed positive loading in PC5 while other traits showed negative loadings. The biplot diagram (Fig. 2) between PC1 and PC2 in this study describes the distribution and type of diversity for yield attributing traits. The maximum vector length was displayed by the number of total tillers, suggesting that it contributed to the overall divergence. This was followed by the number of effective tillers and spikelet fertility. The direction of correlation between the traits is indicated by the angle between the trait vectors. A positive correlation is indicated by an acute angle (less than 90 degrees) between vectors, a negative correlation by an obtuse angle (more than 90 degrees) and no correlation by a right angle (90 degrees). Days to 50% flowering, plant height, flag leaf length, panicle length, number of filled grains per panicle and hundred seed weight were among the twelve features examined for yield-attributing traits; these traits had a positive

**Fig. 6:** The Biplot of aromatic rice genotypes for PC1 and PC2 for grain quality traits.

correlation with grain yield per plant. In this case, there is a negative association between the number of unfilled grains per panicle and the amount of grain produced per plant since the angles are more than 90 degrees. The best-performing genotypes for those traits would be those that are found in the same quadrant as the trait vector. Genotypes such as Jeera phool parent, Gopalbhog parent and Lohandi Mutant-9 are closely aligned with traits like grain yield per plant, filled grains, plant height and panicle length indicating their superior performance for yield-related attributes. TCVM and CG Jawaphool Trombay, Nagri Dubraj parent are located in the direction of total tillers and number of effective tillers suggesting higher tillering ability (Fig. 3).

Based on particular components, these features can be utilized to choose among the many genotypes. The PCA results of the present study align well with earlier findings. Traits like filled grains per panicle, plant height and spikelet fertility contributing to PC1 were also reported by Bharathi *et al.*, (2014) and Kumar *et al.*, (2015). Effective and total tillers influencing PC2 were similarly observed by Raut *et al.*, (2017). The role of hundred seed weight and grain yield in PC3 agrees with Roy *et al.*, (2020). Contributions of days to 50% flowering and sterile spikelets in PC4 and spikelet fertility in PC5, were also consistent with findings by Sahu *et al.*, (2022) and Yadav *et al.*, (2018), respectively.

For grain quality characters 5 PCs found containing 84.446% total variability among the characters studied. Among the five principal components of yield attributing traits, PC1 shared high proportion of total variation 28.078% and the rest of the four principal components viz., PC2, PC3, PC4 and PC5 contributed 18.541%, 17.975%, 11.913% and 7.939% of the total variance respectively. Scree plot elucidated the variation percentage between Eigenvalues and the Principal components (Fig. 4). The proportion of variance, cumulative proportion and eigen values are given in Table 4.

The PC1 was mostly influenced by the paddy length (0.338), kernel length (0.333), kernel L/B (0.437), cooked rice length (0.032) and negatively influenced by the all-other characters. The PC2 was mostly influenced by the milling % (0.109), hulling % (0.041), kernel L/B (0.097), amylose content (0.113), head rice recovery % (0.108). The PC3 was mostly influenced by the elongation ratio (0.3), cooked rice width (0.117), kernel width (0.096). The PC4 was mostly influenced by the kernel L/B (0.225), cooked rice length (0.572), cooked rice width (0.488), elongation ratio (0.471). The PC5 was mostly influenced by paddy width (0.201), kernel width (0.103), cooked rice

width (0.302), head rice recovery % (0.146). It is clearly visible in the graph that the maximum variation was observed in PC1, PC2, PC3, PC4 and PC5.

The biplot diagram (Fig. 5) between PC1 and PC2 explained the distribution and the nature of diversity for variables. The kernel length displayed the longest vector length, signifying its impact on the overall divergence. This was followed by the kernel width, paddy width and kernel L/B. Head rice recovery percentage was positively correlated with the hulling percentage, milling percentage, kernel width and paddy width. Because of the greater than 90-degree angle between them, the head rice recovery in this case was negatively correlated with kernel length, paddy width and kernel L/B. For such traits, the best performing genotypes would be those that are found in the same quadrant as the trait vector. Genotypes like Alsakar parent, Kumbhdaphool Mutant-1 and Kumbhdaphool parent are positively associated with kernel width, paddy width and amylose content. CG Jawaphool Trombay and Jawaphool parent align with hulling % and milling % (Fig. 6).

The present PCA findings on grain quality traits are consistent with earlier studies. Similar contributions of kernel L/B ratio, cooked rice dimensions and elongation ratio to principal components were reported by Roy *et al.*, (2020) and Ravi *et al.*, (2020). Kumar *et al.*, (2015) also observed significant loading of amylose content, milling percentage and head rice recovery in the early PCs, confirming their importance in grain quality-based genetic divergence among rice genotypes.

Acknowledgement

The research was conducted as part of a M. Sc. research work titled "Assessment of genetic diversity using PCA for yield attributing and grain quality traits in aromatic short-grain rice (*Oryza sativa* L.) mutants along with their parental lines" by the primary author at the College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Raipur.

References

- Anderson, T.W. (1972). *An introduction to multivariate statistical analysis* (2nd ed.). Wiley.
- FAO (2024). *World rice market and trade report 2024-25*. Food and Agriculture Organization of the United Nations.
- Ghareyazie, B., Xu J. and Bennett J. (1997). Genetic analysis of aroma in rice. *Rice Genetics Newsletter*, **14**, 20-22.
- Gri. Stat. (2024). *State-wise agricultural statistics - Rice productivity and area trends*. Directorate of Economics and Statistics, Ministry of Agriculture, India.
- IRRI (2023). *Annual report 2023*. International Rice Research Institute.

- Koutroubas, S.D., Mazzini F., Pons B. and Ntanos D.A. (2004). Grain quality variation and relationships with morphophysiological traits in rice. *European Journal of Agronomy*, **21**(3), 285-293.
- Maji, A.T. and Shaibu A.A. (2012). Application of principal component analysis to morphological characterization of rice (*Oryza sativa* L.) germplasm. *Cereal Research Communications*, **40**(3), 439-449.
- Ministry of Agriculture and Farmers Welfare (2025). *Fourth advance estimates of agricultural production 2024-25*. Government of India.
- Nachimuthu, V.V., Robin S., Rajeswari S., Manonmani S., Chandrababu R., Raveendran M. and Samiyappan R. (2014). Evaluation of rice germplasm and genetic diversity analysis using morphological and SSR markers. *Electronic Journal of Plant Breeding*, **5**(3), 538-543.
- Nachimuthu, V.V., Robin S., Sudhakar D., Raveendran M., Rajeswari S. and Raveendran T.S. (2014). Evaluation of rice genetic diversity and variability using multivariate analysis. *Electronic Journal of Plant Breeding*, **10**(3), 1095-1104.
- Oyelola, B.A. (2012). Statistical approaches to analyzing diversity in crop germplasm: A review of multivariate tools. *International Journal of Agricultural Biosciences*, **1**(1), 16-20.
- Ramakrishnan, C., Thirugnanakumar S., Mohan R.J. and Raveendran M. (2016). Principal component analysis of rice (*Oryza sativa* L.) germplasm accessions for yield and quality traits. *International Journal of Agricultural Sciences*, **8**(5), 1260-1263.
- Sivaranjani, A.K.P., Anandan A., Eswaran R. and Prakash M. (2010). Genetic divergence analysis for grain yield and its component traits in traditional and improved rice genotypes. *Electronic Journal of Plant Breeding*, **1**(4), 512-516.
- Sohgaura, R., Srivastava R., Dwivedi S., Yadav R. and Kumar S. (2015). Genetic evaluation of quality traits in indigenous rice landraces. *Oryza*, **52**(4), 345-351.
- Vanaja, T. and Babu L.C. (2006). Genetic variability and diversity studies in rice. *Crop Research*, **32**(1), 63-65.
- Yumnam, J., Shamsudheen M., Ahmed S. and Immanuel S. (2015). Genetic diversity and variability analysis in aromatic rice genotypes. *Oryza*, **52**(1), 70-75.